

SEQUENCE LISTING

(1) GENERAL INFORMATION

(i) APPLICANT: Chatterjee, Subroto

(ii) TITLE OF THE INVENTION: RECOMBINANT N-SMASES AND NUCLEIC ACIDS
ENCODING SAME

(iii) NUMBER OF SEQUENCES: (7)

(iv) CORRESPONDENCE ADDRESS:

- (A) ADDRESSEE: Dike, Bronstein, Roberts & Cushman, LLP
- (B) STREET: 130 Water Street
- (C) CITY: Boston
- (D) STATE: MA
- (E) COUNTRY: USA
- (F) ZIP: 02109

(v) COMPUTER READABLE FORM:

- (A) MEDIUM TYPE: Diskette
- (B) COMPUTER: IBM Compatible
- (C) OPERATING SYSTEM: DOS
- (D) SOFTWARE: FastSEQ Version 1.5

(vi) CURRENT APPLICATION DATA:

- (A) APPLICATION NUMBER:
- (B) FILING DATE:
- (C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:

- (A) APPLICATION NUMBER:
- (B) FILING DATE:

(viii) ATTORNEY/AGENT INFORMATION:

- (A) NAME: Corless, Peter F
- (B) REGISTRATION NUMBER: 33,860
- (C) REFERENCE/DOCKET NUMBER: 46906

(ix) TELECOMMUNICATION INFORMATION:

- (A) TELEPHONE: 617-523-3400
- (B) TELEFAX: 617-523-6440
- (C) TELEX:

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1197 base pairs
- (B) TYPE: nucleic acid

657220.622220

- (ii) MOLECULE TYPE: cDNA
- (iii) HYPOTHETICAL: NO
- (iv) ANTISENSE: NO
- (v) FRAGMENT TYPE:
- (vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

ATGATGACAT	ATCACGAAAC	GCGCGCGTTG	GCTCAAAGCG	ACTTACAGCA	ACTCTATGCG	60
GCACTTGAAA	CAACTGAATT	TGGCGCTTAC	TTTGCGACAC	CCGCTGATGA	TACTTTACGT	120
TTTGGCATTG	GCGCAATCGC	TACGGCAAAA	ACGGCTCAGG	CATTACAAGG	TGCGGTTGTT	180
TTTGGTGCGC	AGTCATTTGA	TGAACAAGAG	TACCCGCACT	CTGAATTGAT	GGCGGGTTTT	240
TGGTTTGTCC	CCGAAGTGAT	GGTGACCATC	GCGGCAGATA	AAATCACGTT	CGGATCAGAT	300
ACCGTATCTG	ATTTTACGAC	GTGGCTGGCG	CAGTTCGTGC	CAAAACAGCC	AAATACGGTG	360
ACCACTAGTC	ATGTGACAGA	TGAAGTGGAT	TGGATCGAAC	GGACAGAGAA	TTTGATTGAT	420
ACCTTAGCCA	TCGATCAAAC	CTTAGCCAAA	GTCGTTTTTG	GTCGGCAACA	GACCCTGCAG	480
TTATCCGACA	CGTTACGACT	GGCACAAATT	ATTCGTGCGT	TAGCTGAGCA	GGCGAATACG	540
TATCATGTGG	TTTTAAAGCG	ACATGATGAA	TTGTTTATTT	CAGCAACACC	GGAACGGTTA	600
GTGGCTATGT	CAGGTGGTCA	GATCGCTACG	GCGGCGGTCT	CTGGGACAAG	CCGGCGCGGG	660
ACGGATGGCG	CTGACGATAT	CGCGTTAGGC	GAAGCGTTGT	TAGCCAGTCA	GAAAAACCGC	720
GATTGAACATC	AATATGTCGT	GGCAAGTATC	ACGACACGCT	TGCAAGACGT	GACGACGTCG	780
CTTAAAGGTGC	CGGCCATGCC	AAGTTTACTC	AAAAATAAGC	AAGTTCAGCA	TTTGTACACA	840
CCCAATTACAG	GGGACATTGC	GGCACATTTA	AGTGTGACCG	CGATTGTTGA	CCGCTTGCA	900
CCCAACACCAG	CCTGCGGTGG	CGTCCACGCT	GAAGCGGCCC	TGTATTACAT	TGCGACCCAT	960
GAGAAGACAC	CTCGTGGCTT	GTTTGCAGGT	CCTATTGGCT	ATTTTACCGC	AGATAATAGT	1020
GGGGAATTTG	TGGTTGGCAT	CCGTTCCATG	TATGTGAATC	AAACGCAGCG	ACGAGCAACT	1080
TTATTTGCTG	GTGCCGGGAT	TGTGGCTGAC	TCCGATGCGC	AACAAGAATA	TGAAGAACT	1140
GGGTTGAAAT	TTGAACCCAT	GCGGCAATTG	TTAAAGGACT	ACAATCATGT	CGAATGA	1197

(2) INFORMATION FOR SEQ ID NO:2:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 397 amino acids
 - (B) TYPE: amino acid
- (ii) MOLECULE TYPE: protein
- (iii) HYPOTHETICAL: NO
- (iv) ANTISENSE: NO
- (v) FRAGMENT TYPE: N-terminal
- (vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met	Met	Thr	Tyr	His	Glu	Thr	Arg	Ala	Leu	Ala	Gln	Ser	Asp	Leu	Gln
1				5					10					15	
Gln	Leu	Tyr	Ala	Ala	Leu	Glu	Thr	Thr	Glu	Phe	Gly	Ala	Tyr	Phe	Ala
			20					25					30		
Thr	Pro	Ala	Asp	Asp	Thr	Leu	Arg	Phe	Gly	Ile	Gly	Ala	Ile	Ala	Thr
		35					40				45				
Ala	Lys	Thr	Ala	Gln	Ala	Leu	Gln	Gly	Ala	Val	Phe	Gly	Ala	Gln	Ser
	50					55				60					
Phe	Asp	Glu	Gln	Glu	Tyr	Pro	Gln	Ser	Glu	Leu	Met	Ala	Gly	Phe	Trp

65 70 75 80
Phe Val Pro Glu Val Met Val Thr Ile Ala Ala Asp Lys Ile Thr Phe
85 90 95
Gly Ser Asp Thr Val Ser Asp Phe Thr Trp Leu Ala Gln Phe Val
100 105 110
Pro Lys Gln Pro Asn Thr Val Thr Ser His Val Thr Asp Glu Val
115 120 125
Asp Trp Ile Glu Arg Thr Glu Asn Leu Ile Asp Thr Leu Ala Ile Asp
130 135 140
Gln Thr Leu Ala Lys Val Val Phe Gly Arg Gln Gln Thr Leu Gln Leu
145 150 155 160
Ser Asp Thr Leu Arg Leu Ala Gln Ile Ile Arg Ala Leu Ala Glu Gln
165 170 175
Ala Asn Thr Tyr His Val Val Leu Lys Arg His Asp Glu Leu Phe Ile
180 185 190
Ser Ala Thr Pro Glu Arg Leu Val Ala Met Ser Gly Gly Gln Ile Ala
195 200 205
Thr Ala Ala Val Ala Gly Thr Ser Arg Arg Gly Thr Asp Gly Ala Asp
210 215 220
Asp Ile Ala Leu Gly Glu Ala Leu Leu Ala Ser Gln Lys Asn Arg Ile
225 230 235 240
Glu His Gln Tyr Val Val Ala Ser Ile Thr Thr Arg Leu Gln Asp Val
245 250 255
Thr Thr Ser Leu Lys Val Pro Ala Met Pro Ser Leu Leu Lys Asn Lys
260 265 270
Gln Val Gln His Leu Tyr Thr Pro Ile Thr Gly Asp Ile Ala Ala His
275 280 285
Leu Ser Val Thr Ala Ile Val Asp Arg Leu His Pro Thr Pro Ala Leu
290 295 300
Gly Gly Val Pro Arg Glu Ala Ala Leu Tyr Tyr Ile Ala Thr His Glu
305 310 315 320
Lys Thr Pro Arg Gly Leu Phe Ala Gly Pro Ile Gly Tyr Phe Thr Ala
325 330 335
Asp Asn Ser Gly Glu Phe Val Val Gly Ile Arg Ser Met Tyr Val Asn
340 345 350
Gln Thr Gln Arg Arg Ala Thr Leu Phe Ala Gly Ala Gly Ile Val Ala
355 360 365
Asp Ser Asp Ala Gln Gln Glu Tyr Glu Glu Thr Gly Leu Lys Phe Glu
370 375 380
Pro Met Arg Gln Leu Leu Lys Asp Tyr Asn His Val Glu
385 390 395

(2) INFORMATION FOR SEQ ID NO:3:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 7 amino acids
 - (B) TYPE: amino acid
- (ii) MOLECULE TYPE: protein
- (iii) HYPOTHETICAL: NO
- (iv) ANTISENSE: NO
- (v) FRAGMENT TYPE: N-terminal

(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

Thr Ser Leu Lys Val Pro Ala
1 5

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 7 amino acids
- (B) TYPE: amino acid
- (ii) MOLECULE TYPE: protein
- (iii) HYPOTHETICAL: NO
- (iv) ANTISENSE: NO
- (v) FRAGMENT TYPE: N-terminal
- (vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Arg Ser Ile Thr Val Arg Val
1 5

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18 base pairs
- (B) TYPE: nucleic acid
- (ii) MOLECULE TYPE: cDNA
- (iii) HYPOTHETICAL: NO
- (iv) ANTISENSE: NO
- (v) FRAGMENT TYPE:
- (vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

TTGCGGCACT ATTAGGTG

18

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18 base pairs
- (B) TYPE: nucleic acid
- (ii) MOLECULE TYPE:
- (iii) HYPOTHETICAL: NO
- (iv) ANTISENSE: NO
- (v) FRAGMENT TYPE:
- (vi) ORIGINAL SOURCE:

09682379.033499

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

CGCCAAATGCC AAAACGTA

18

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 50 base pairs

(B) TYPE: nucleic acid

(ii) MOLECULE TYPE:

(iii) HYPOTHETICAL: NO

(iv) ANTISENSE: NO

(v) FRAGMENT TYPE:

(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

GATCCATGAT GACATATCAC GAAACGCGCG TTTCGTGATA TGTCATCATG

50

0928279-033199